VERJETNOSTNE PORAZDELITVE V PROGRAMU MATHEMATICA

Dejan Zupan

Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo

Osnovni ukazi

- gostota $f_X(x)$:

  PDF[porazdelitev, x]

- porazdelitvena funkcija $F_X(x)$:

  CDF[porazdelitev, x]

- verjetnost $P[X \leq x]$:

  Probability[X ≤ x, X ≈ porazdelitev]

- srednja vrednost $E[X]$:

  Mean[porazdelitev]

- varianca $var[X]$:

  Variance[porazdelitev]

Porazdelitve

- veliko število vgrajenih porazdelitev:

$$\mathbf{NameDistribution[p1, p2, ...]}$$

- imena poznamo iz teorije verjetnostnega računa in statistike
- število in pomen parametrov je odvisno od tipa porazdelitve
- pripravimo lahko tudi nove porazdelitve
- transformacija slučajne spremenljivke $g(X)$:

$$\mathbf{Nova = TransformedDistribution[g(X), X \approx porazdelitev]}$$

- transformacija slučajnega vektorja $g(X, Y)$:

$$\mathbf{Nova = TransformedDistribution[g(X, Y), \{X \approx porazdelitev, Y \approx porazdelitev\}]}$$

Porazdelitve

- Binomska **BinomialDistribution**$[\mathbf{n}, \mathbf{p}]$
- Poissonova **PoissonDistribution**$[\nu]$
- Normalna **NormalDistribution**$[\mathbf{m}, \sigma]$
- Standardizirana normalna **NormalDistribution**$[\,]$
- Logaritemsko normalna **LogNormalDistribution**$[\mathbf{m}, \sigma]$
- Gumbelova za minimume (ekstremnih vrednosti I) **GumbelDistribution**$[\mathbf{v}, 1/\lambda]$
- Frechetova za maksimume (ekstremnih vrednosti II) **FrechetDistribution**$[\mathbf{k}, \mathbf{u}-\epsilon, \epsilon]$
- Weibullova za minimume (ekstremnih vrednosti III) **WeibullDistribution**$[\mathbf{k}, \mathbf{v}-\epsilon, \epsilon]$

Ocenjevanje parametrov

- točkovne ocene parametrov:

  FindDistributionParameters[podatki, porazdelitev]
  porazdelitevO = EstimatedDistribution[podatki, porazdelitev]

- podatki so rezultat vzorčenja:

  podatki $= \{x_1, x_2, \ldots, x_N\}$

- porazdelitev je poljubna vgrajena ali pripravljena porazdelitev
- izbira metode:

  ParameterEstimator → "ImeMetode"

- privzeta metoda je metoda največjega verjetja:

  "MaximumLikehood"

Metode točkovnih ocen

- metoda največjega verjetja $\max\left(\sum_{i=1}^{n} \ln p_X\left(x_i, \boldsymbol{a}\right)\right)$:

  "MaximumLikelihood"

- metoda momentov $\widehat{m}_i = m^{(i)}\left(\boldsymbol{a}\right)$ za $i = 1, \ldots, N_a$:

  "MethodOfMoments"

- metoda centralnih momentov $\widehat{\mu}_i = \mu^{(i)}\left(\boldsymbol{a}\right)$ za $i = 1, \ldots, N_a$:

  "MethodOfCentralMoments"

- kombinacija metod:

  Solve[{Mean[podatki] == Mean[porazdelitev[{a, b}]],
  Variance[podatki] == Variance[porazdelitev[a, b]]}, {a, b}]

Statistike vzorca

- povprečje $\bar{X}$:

  Mean[podatki]

- varianca S_X^{*2}:

  Variance[podatki]

- r-ti moment $\frac{1}{n}\sum x_i^r$:

  Moment[podatki, r]

- r-ti centralni moment $\frac{1}{n}\sum (x_i - \bar{x})^r$:

  CentralMoment[podatki, r]

Intervalne ocene

- interval zaupanja za pričakovano vrednost

$$\mathbf{Needs}[\text{"}\mathbf{HypothesisTesting}\text{`"}]$$
$$\mathbf{MeanCI}[\mathbf{podatki}]$$

- stopnja zaupanja:

$$\mathbf{ConfidenceLevel} \to 1 - \alpha$$

- brez dodatnega paketa:

$$\mathbf{Mean}[\mathbf{podatki}] \pm \sqrt{\frac{[\mathbf{Variance}[\mathbf{podatki}]}{\mathbf{Length}[\mathbf{podatki}]}}$$
$$\mathbf{InverseCDF}\left[\mathbf{StudentTDistribution}[\mathbf{Length}[\mathbf{podatki}] - 1], 1 - \frac{\alpha}{2}\right]$$

Intervalne ocene

- interval zaupanja za varianco:

$$\mathbf{Needs[\text{"HypothesisTesting`"}]}$$
$$\mathbf{VarianceCI[podatki]}$$

- brez dodatnega paketa:

$$\frac{\mathbf{Variance[podatki](Length[podatki]-1)}}{\mathbf{InverseCDF\left[ChiSquareDistribution[Length[podatki]-1],1-\frac{\alpha}{2}\right]}}$$

$$\frac{\mathbf{Variance[podatki](Length[podatki]-1)}}{\mathbf{InverseCDF\left[ChiSquareDistribution[Length[podatki]-1],\frac{\alpha}{2}\right]}}$$